

Quantitative statistical stability, speed of convergence to equilibrium and partially hyperbolic skew products.

Stefano Galatolo *

April 25, 2017

Abstract

We consider a general relation between fixed point stability of suitably perturbed transfer operators and convergence to equilibrium (a notion which is strictly related to decay of correlations). We apply this relation to deterministic perturbations of a class of (piecewise) partially hyperbolic skew products whose behavior on the preserved fibration is dominated by the expansion of the base map. In particular we apply the results to power law mixing toral extensions. It turns out that in this case, the dependence of the physical measure on small deterministic perturbations, in a suitable anisotropic metric is at least Hölder continuous, with an exponent which is explicitly estimated depending on the arithmetical properties of the system. We show explicit examples of toral extensions having actually Hölder stability and non differentiable dependence of the physical measure on perturbations.

1 Introduction

The statistical stability of dynamical systems is well understood in the uniformly hyperbolic case. In this case quantitative estimates are available, proving the Lipschitz and even differentiable dependence of the physical measure under perturbations (see e.g [3], [7] or [28] and related references, for recent surveys where also some result beyond the uniformly expanding case are discussed). For systems having a non uniformly hyperbolic behavior, and in presence of discontinuities, the situation is more complicated and much less is known. Qualitative results and some quantitative ones (providing precise information on the modulus of continuity) are known under different assumptions or in families of cases, and there is not yet a general understanding of the statistical stability in that cases (see. e.g. [1], [2], [5], [8], [9], [10], [11], [16], [18], [19], [25], [32], [33]).

*Dipartimento di Matematica, Università di Pisa, Via Buonarroti 1, Pisa - Italy. Email: galatolo@dm.unipi.it

In this paper we use a general quantitative relation between speed of convergence to equilibrium and statistical stability (applied in [19] to maps with indifferent fixed points) giving nontrivial information for systems having different speed of convergence to equilibrium. We show the flexibility of this approach showing a general quantitative stability statement for perturbations of a class of partially hyperbolic skew products having some discontinuities.

The statement is then applied to piecewise constant toral extensions: maps of the kind (X, F) where $X = [0, 1] \times \mathcal{T}^d$, \mathcal{T}^d is the d dimensional torus and $F : X \rightarrow X$ is defined by

$$F(\omega, t) = (T\omega, t + \tau(\omega)) \quad (1)$$

where T is expanding and $\tau : [0, 1] \rightarrow \mathcal{T}^d$ is a piecewise constant function. The qualitative ergodic theory of this kind of systems was studied in several papers (see e.g. [13],[14]). Quantitative results appeared more recently ([17], [15],[29]), proving from different points of views that the speed of correlation decay is generically fast (exponential), but in some cases where τ is piecewise constant, this follows a power law (see [21] or Section 6.1).

We apply our general result to deterministic perturbations of these partially hyperbolic, slowly mixing, discontinuous maps, showing that the physical measure of those systems varies at least Hölder continuously in a kind of anisotropic metric on the space of measures on X^1 . We remark that for a class of smooth toral extensions with fast decay of correlations, a differentiable dependence statement was proved in [18]. We finally show examples of piecewise constant toral extensions where the physical measure of the system actually varies in a Hölder way (and hence not in a differentiable way) with an exponent depending on the arithmetical properties of the system. This shows that in some sense, the general result mentioned above is sharp. We remark that recently, in [33] examples of C^r families of *mostly contracting* diffeomorphisms with strictly Hölder behavior have been given (see also [16] for previous results on Hölder stability of these kinds of partially hyperbolic maps).

The paper is structured as follows: in Section 2 we show a general result estimating quantitatively the stability of operator's fixed points under certain perturbations. The statement is suitable to be applied in estimating the stability of physical measures. In Section 3 we introduce a class of anisotropic spaces adapted to skew products and show some basic properties that make them work quite like L^1 and Bounded Variation real functions spaces in the classical theory for the statistical properties of one dimensional dynamics. The regularity of these spaces is defined by disintegrating along the stable foliation and considering the regularity of the disintegration. In Section 4 we prove a kind of Lasota Yorke inequality. This will be used together with a kind of Helly selection principle proved in Section 3 to estimate the regularity of the invariant measures (like in the classical one dimensional, piecewise expanding case). In

¹We remark that the perturbations allowed on the toral extension systems are quite general. In particular they allow discontinuities, and the invariant measure to become singular with respect to the Lebesgue measure.

Section 5 we consider a class of perturbations of our skew products such that the related transfer operators are near in some sense when applied to (regular) measures and state a first general statement on the statistical stability of such skew products. In Section 6 we introduce the class of slowly mixing, partially hyperbolic toral extensions, to which we apply our general results. In Section 6.2 we deduce from the decay of correlations with Lipschitz observables, a convergence to equilibrium statement adapted to the spaces we consider. In the final part of the section we use these estimates, applying the general statement of Section 2 to our perturbed skew products. Finally in Section 6.3 we show examples of piecewise constant toral extensions where a perturbation of the map of size δ result in a change of physical invariant measure of the size of order δ^β , where $\beta \leq 1$ depends on the Diophantine properties of the map.

Acknowledgement 1 *The work was partially supported by EU Marie-Curie IRSES Brazilian-European partnership in Dynamical Systems (FP7-PEOPLE-2012-IRSES 318999 BREUDS), and by The Leverhulme Trust through Network Grant IN-2014-021.*

2 Quantitative fixed point stability and convergence to equilibrium.

Let us consider a dynamical system (X, T) where X is a metric space and the space $SM(X)$ of Borel measures with sign on X . The dynamics T naturally induces a function $L : SM(X) \rightarrow SM(X)$ which is linear and is called transfer operator. If $\nu \in SM(X)$ then $L[\nu] \in SM(X)$ is defined by

$$L[\nu](B) = \nu(T^{-1}(B))$$

for every measurable set B . If X is a manifold, the measure is absolutely continuous ($d\nu = f \, dm$, where m represents the Lebesgue measure) and T is nonsingular, the operator induces another operator $\tilde{L} : L^1(m) \rightarrow L^1(m)$ acting on measure densities ($\tilde{L}f = \frac{d(L(f \, m))}{dm}$). By a small abuse of notation we will still indicate by L this operator.

An invariant measure is a fixed point for the transfer operator. Let us now see a quantitative stability statement for these fixed points under suitable perturbations of the operator. Let us consider a certain system having a transfer operator L_0 for which we know the speed of convergence to equilibrium (see (2) below). Consider a "nearby" system L_1 having suitable properties: suppose there are two normed vector spaces of measures with sign $B_s \subseteq B_w \subseteq SM(X)$ (the strong and weak space) with norms $\|\cdot\|_w \leq \|\cdot\|_s$ and suppose the operators L_0 and L_1 preserve the spaces: $L_i(B_s) \subseteq B_s$ and $L_i(B_w) \subseteq B_w$ with $i \in \{0, 1\}$. Let us consider

$$V_s := \{f \in B_s, f(X) = 0\}$$

the space of zero average measures in B_s . The speed of convergence to equilibrium of a system will be measured by the speed of contraction to 0 of this space by the iterations of the transfer operator.

Definition 2 Let $\phi(n)$ be a real sequence converging to zero. We say that the system has convergence to equilibrium with respect to norms $\|\cdot\|_w$, $\|\cdot\|_s$ and speed ϕ if $\forall g \in V_s$

$$\|L_0^n(g)\|_w \leq \phi(n)\|g\|_s. \quad (2)$$

Suppose $f_0, f_1 \in B_s$ are fixed probability measures of L_0 and L_1 . The following statement relates the distance between f_0 and f_1 with the distance between L_0 and L_1 and the speed of convergence to equilibrium of L_0 . The proof is elementary, we include it for completeness. Similar quantitative stability statements are used in [22], [20] and [19] to support rigorous computation of invariant measures, get quantitative estimations for the statistical stability of Lorenz like maps and intermittent systems.

Theorem 3 Suppose we have estimations on the following aspects of the operators L_0 and L_1 :

1. (speed of convergence to equilibrium) there is $\phi \in C^0(\mathbb{R})$, $\phi(t)$ decreasing to 0 as $t \rightarrow \infty$ such that L_0 has convergence to equilibrium with respect to norms $\|\cdot\|_w$, $\|\cdot\|_s$ and speed ϕ .
2. (control on the norms of the invariant measures) There is $\tilde{M} \geq 0$ such that

$$\max(\|f_1\|_s, \|f_0\|_s) \leq \tilde{M};$$

3. (iterates of the transfer operator are bounded for the weak norm) there is $\tilde{C} \geq 0$ such that for each n ,

$$\|L_0^n\|_{B_w \rightarrow B_w} \leq \tilde{C}.$$

4. (control on the size of perturbation in the strong-weak norm) Denote

$$\sup_{\|f\|_s \leq 1} \|(L_1 - L_0)f\|_w := \epsilon$$

consider the decreasing function ψ defined as $\psi(x) = \frac{\phi(x)}{x}$, then we have the following explicit estimate

$$\|f_1 - f_0\|_w \leq 2M\tilde{C}\epsilon(\psi^{-1}(\frac{\epsilon\tilde{C}}{2}) + 1).$$

Proof. The proof is a direct computation from the assumptions

$$\begin{aligned} \|f_1 - f_0\|_w &\leq \|L_1^N f_1 - L_0^N f_0\|_w \\ &\leq \|L_1^N f_1 - L_0^N f_1\|_w + \|L_0^N f_1 - L_0^N f_0\|_w \\ &\leq \|L_0^N(f_1 - f_0)\|_w + \|L_1^N f_1 - L_0^N f_1\|_w. \end{aligned}$$

Since $f_1 - f_0 \in V_s$, $\|f_1 - f_0\|_s \leq 2M$,

$$\|f_1 - f_0\|_w \leq 2M\phi(n) + \|L_1^N f_1 - L_0^N f_1\|_w$$

but

$$(L_0^N - L_1^N) = \sum_{k=1}^N L_0^{N-k} (L_0 - L_1) L_1^{k-1}$$

hence

$$\begin{aligned} -(L_1^N - L_0^N) f_1 &= \sum_{k=1}^N L_0^{N-k} (L_0 - L_1) L_1^{k-1} f_1 \\ &= \sum_{k=1}^N L_0^{N-k} (L_0 - L_1) f_1 \end{aligned}$$

then

$$\|f_1 - f_0\|_w \leq 2M\phi(N) + \epsilon MN\tilde{C}.$$

Now consider the function ψ defined as $\psi(x) = \frac{\phi(x)}{x}$, choose N such that $\psi^{-1}(\frac{\epsilon\tilde{C}}{2}) \leq N \leq \psi^{-1}(\frac{\epsilon\tilde{C}}{2}) + 1$, in this way $\frac{\phi(N)}{N} \leq \frac{\epsilon\tilde{C}}{2} \leq \frac{\phi(N-1)}{N-1}$ and

$$\|f_1 - f_0\|_w \leq 2M\tilde{C}\epsilon(\psi^{-1}(\frac{\epsilon\tilde{C}}{2}) + 1).$$

■

Remark 4 We remark that if $\phi \rightarrow 0$ then also $\|f_1 - f_0\|_w \xrightarrow{\epsilon \rightarrow 0} 0$, and the system is strongly statistically stable. If $\phi(x) = Cx^{-\alpha}$ then $\psi(x) = Cx^{-\alpha-1}$ and $\epsilon(\psi^{-1}(\epsilon) + 1) \sim \epsilon^{1-\frac{1}{\alpha+1}}$ and we have the estimate for the modulus of continuity

$$\|f_1 - f_0\|_w \leq K_1 \epsilon^{1-\frac{1}{\alpha+1}}$$

where the constant K_1 depends on M, \tilde{C}, C and not on the distance between the operators measured by ϵ .

3 Spaces we consider

Our approach is based on the study of the transfer operator restricted to a suitable space of measures with sign. We introduce a space of regular measures where we can find the invariant measure of our systems, and the ones of suitable perturbations of it. We hence consider some measure spaces adapted to skew products. The approach is taken from [20] (see also [4]) where it was used for the Lorenz-like two dimensional maps. Let us consider a map $F : X \rightarrow X$ where $X = [0, 1] \times M$, and M is a compact manifold with boundary, such that

$$F(x, y) = (T(x), G(x, y)). \quad (3)$$

Suppose F satisfies the following conditions:

Sk1 Suppose T is $\frac{1}{\lambda}$ -expanding² and it has $C^{1+\xi}$ branches³ which are onto. The branches will be denoted by T_i , $i \in [1, \dots, q]$.

Sk2 Consider the F -invariant foliation $\mathcal{F}^s := \{\{x\} \times M\}_{x \in [0,1]}$. We suppose that the behavior on \mathcal{F}^s is dominated by λ : there is $\alpha \in \mathbb{R}$ with $\lambda^\xi \alpha < 1$, such that for all $x \in [0, 1]$ holds

$$|G(x, y_1) - G(x, y_2)| \leq \alpha |y_1 - y_2| \quad \text{for all } y_1, y_2 \in M. \quad (4)$$

Sk3 For each $p \leq \xi$ there is $A > 0$, such that

$$\hat{H} := \sup_{r \leq A} \frac{1}{r^p} \int \sup_{y \in M, x_1, x_2 \in B(x, r)} |G(x_1, y) - G(x_2, y)| dx < \infty$$

Remark 5 We remark that Sk3 allows discontinuities in G , provided a kind of bounded variation regularity is respected. Sk2 allows a dominated expansion in the fibers direction. Furthermore, by Sk1 the transfer operator of the map T satisfies a Lasota Yorke inequality of the kind

$$\|L_T^n(\mu)\|_{BV} \leq A_T \lambda^n \|\mu\|_{BV} + B_T \|\mu\|_1 \quad (5)$$

where $\|\mu\|_{BV}$ is the generalized bounded variation norm (see [26]); for some constant A_T and B_T depending on the map.

Definition 6 We say that a family of maps $F_\delta = (T_\delta(x), G_\delta(x, y))$ satisfies Sk1, ..., Sk3 uniformly, if each T_δ is piecewise expanding, with onto $C^{1+\xi}$ branches, admitting a uniform expansion rate $\frac{1}{\lambda}$, a uniform α , a uniform Hölder constant C_h , a uniform second coefficient of the Lasota Yorke inequality B_{T_δ} and furthermore the family G_δ satisfies Sk3 with a uniform bound on the constant \hat{H} .

We construct now some function spaces which are suitable for the systems we consider. The idea is to consider spaces of measures with sign, with suitable norms constructed by disintegrating measures along the central foliation. In this way a measure on X will be seen as a collection (a path) of measures on the leaves. In the central direction (on the leaves) we will consider a norm which is the dual of the Lipschitz norm. In the expanding direction we will consider the L^1 norm and a suitable variation norm. These ideas will be implemented in the next paragraphs.

Let (X, d) be a compact metric space, $g : X \rightarrow \mathbb{R}$ be a Lipschitz function and let $Lip(g)$ be its best Lipschitz constant, i.e.

$$Lip(g) = \sup_{x, y \in X} \left\{ \frac{|g(x) - g(y)|}{d(x, y)} \right\}.$$

²We suppose that $\inf_{x \in [0,1]} T'(x) \geq \frac{1}{\lambda}$ for some $\lambda < 1$.

³More precisely we suppose that there are $\xi, C_h \geq 0$ such that

$$\frac{1}{|T'_i \circ T_i^{-1}(\gamma_2)|} - \frac{1}{|T'_i \circ T_i^{-1}(\gamma_1)|} \leq C_h d(\gamma_1, \gamma_2)^\xi, \forall \gamma_1, \gamma_2 \in [0, 1].$$

Definition 7 Given two signed Borel measures μ and ν on X , we define a **Wasserstein-Kantorovich Like** distance between μ and ν by

$$W_1(\mu, \nu) = \sup_{Lip(g) \leq 1, \|g\|_\infty \leq 1} \left| \int g d\mu - \int g d\nu \right|. \quad (6)$$

Let us denote

$$\|\mu\|_{W_1} := W_1(0, \mu).$$

As a matter of fact, $\|\cdot\|_{W_1}$ defines a norm on the vector space of signed measures defined on a compact metric space.

Let $\mathcal{SB}(\Sigma)$ be the space of Borel signed measures on Σ . Given $\mu \in \mathcal{SB}(\Sigma)$ denote by μ^+ and μ^- the positive and the negative parts of it ($\mu = \mu^+ - \mu^-$).

Denote by \mathcal{AB} the set of signed measures $\mu \in \mathcal{SB}(\Sigma)$ such that its associated marginal signed measures, $\mu_x^\pm = \pi_x^* \mu^\pm$ are absolutely continuous with respect to the Lebesgue measure m , on $[0, 1]$ i.e.

$$\mathcal{AB} = \{\mu \in \mathcal{SB}(\Sigma) : \pi_x^* \mu^+ \ll m \text{ and } \pi_x^* \mu^- \ll m\}$$

where $\pi_x : X \rightarrow [0, 1]$ is the projection defined by $\pi_x(x, y) = x$ and π_x^* is the associated pushforward map.

Let us consider a finite positive measure $\mu \in \mathcal{AB}$ on the space X foliated by the preserved leaves $\mathcal{F}^c = \{\gamma_l\}_{l \in N_1}$ such that $\gamma_l = \pi_x^{-1}(l)$. We will also call \mathcal{F}^c as the *central foliation*. Let us denote $\mu_x = \pi_x^* \mu$ and let ϕ_μ be its density ($\mu_x = \phi_\mu m$). The Rokhlin disintegration theorem describes a disintegration of μ by a family $\{\mu_\gamma\}_\gamma$ of probability measures on the central leaves⁴ in a way that the following holds.

Remark 8 The disintegration of a measure μ is the μ_x -unique measurable family $(\{\mu_\gamma\}_\gamma, \phi_\mu)$ such that, for every measurable set $E \subset X$ it holds

$$\mu(E) = \int_{[0,1]} \mu_\gamma(E \cap \gamma) d\mu_x(\gamma). \quad (7)$$

Definition 9 Let $\pi_{\gamma,y} : \gamma \rightarrow M$ be the restriction $\pi_y|_\gamma$, where $\pi_y : X \rightarrow M$ is the projection defined by $\pi_y(x, y) = y$ and $\gamma \in \mathcal{F}^c$. Given a positive measure $\mu \in \mathcal{AB}$ and its disintegration along the stable leaves \mathcal{F}^c , $(\{\mu_\gamma\}_\gamma, \phi_\mu)$, we define the **restriction of μ on γ** as the positive measure $\mu|_\gamma$ on M (not on the leaf γ) defined as

$$\mu|_\gamma = \pi_{\gamma,y}^* (\phi_\mu(\gamma) \mu_\gamma).$$

Definition 10 For a given signed measure $\mu \in \mathcal{AB}$ and its decomposition $\mu = \mu^+ - \mu^-$, define the **restriction of μ on γ** by

$$\mu|_\gamma = \mu^+|_\gamma - \mu^-|_\gamma. \quad (8)$$

⁴In the following to simplify notations, when no confusion is possible we will indicate the generic leaf or its coordinate with γ .

Definition 11 Let $\mathcal{L}^1 \subseteq \mathcal{AB}$ be defined as

$$\mathcal{L}^1 = \left\{ \mu \in \mathcal{AB} : \int_{[0,1]} \|\mu|_\gamma\|_{W_1} dm(\gamma) < \infty \right\} \quad (9)$$

and define norm on it, $\|\cdot\|_1 : \mathcal{L}^1 \rightarrow \mathbb{R}$, by

$$\|\mu\|_1 = \int_{[0,1]} \|\mu|_\gamma\|_{W_1} dm(\gamma). \quad (10)$$

The notation we use for this norm is similar to the usual L^1 norm. Indeed this is formally the case if we associate to μ , by disintegration, a path $G_\mu : [0, 1] \rightarrow (\mathcal{SB}(M), \|\cdot\|_{W_1})$ defined by $G_\mu(\gamma) = \mu|_\gamma$. In this case, this will be the L^1 norm of the path.

3.1 The transfer operator associated to F and basic properties of \mathcal{L}^1

Let us now consider the transfer operator L_F associated with F . Being a push forward map, the same function can be also denoted by F^* we will use this notation sometime. There is a nice characterization of the transfer operator in our case, which makes it work quite like a one dimensional operator. For the proof see [20].

Proposition 12 (Perron-Frobenius like formula) *Let us consider a skew product map F satisfying Sk1 and Sk2. For a given leaf $\gamma \in \mathcal{F}^s$, define the map $F_\gamma : M \rightarrow M$ by*

$$F_\gamma = \pi_y \circ F|_\gamma \circ \pi_{\gamma,y}^{-1}.$$

For all $\mu \in \mathcal{L}^1$ and for almost all $\gamma \in [0, 1]$ it holds

$$(L_F \mu)|_\gamma = \sum_{i=1}^q \frac{F_{T_i^{-1}(\gamma)}^* \mu|_{T_i^{-1}(\gamma)}}{|T_i' \circ T_i^{-1}(\gamma)|}. \quad (11)$$

We recall some results showing that the transfer operator associated to a Lipschitz function is also Lipschitz with the same constant, for the " $\|\cdot\|_1$ " distance, and moreover, that the transfer operator of a map satisfying Sk1, ..., Sk3 is also Lipschitz with the same constant for the " $\|\cdot\|_1$ " norm. In particular, if $\alpha \leq 1$ the transfer operator is a weak contraction, like it happen for the L^1 norm on the one dimensional case (for the proof and more details see [20]).

Lemma 13 *If $G : Y \rightarrow Y$, where Y is a metric space is α -Lipschitz, for every Borel measure with sign μ it holds*

$$\|L_G \mu\|_{W_1} \leq \alpha \|\mu\|_{W_1}.$$

If $\mu \in \mathcal{L}^1$ and $F : [0, 1] \times M \rightarrow [0, 1] \times M$ satisfies Sk1, ..., Sk3 then

$$\|L_F \mu\|_1 \leq \alpha \|\mu\|_1. \quad (12)$$

3.2 The strong norm

We consider a norm which is stronger than the \mathcal{L}^1 norm. The idea is to consider a disintegrated measure as a path of measures on the preserved foliation and define a kind of bounded variation regularity for this path, in a way similar to what was done in [26] for real functions.

For this strong space we will prove a regularization inequality, similar to the Lasota Yorke ones. We will use this inequality to prove the regularity of the invariant measure of the family of skew products we consider.

Let us consider $\mu \in \mathcal{L}^1$. Let us define

$$osc(\mu, x, r) = \operatorname{esssup}_{\gamma_2, \gamma_1 \in B(x, r)} (W_1(\mu|_{\gamma_1}, \mu|_{\gamma_2}))$$

and

$$var_p(\mu, r) = \int_{[0,1]} r^{-p} osc(\mu, x, r) \, dx.$$

Now let us choose $A > 0$ and consider $var_p(\mu) := \sup_{r \leq A} var_p(\mu, r)$. Finally we define $p - BV$ norm as:

$$\|\mu\|_{p-BV} = \|\mu\|_1 + var_p(\mu).$$

Let us consider $1 \geq p \geq 0$ and the following space of measures

$$p - BV = \{\mu \in \mathcal{L}^1, \|\mu\|_{p-BV} < \infty\}.$$

This will play the role of the strong space in the present case.

Remark 14 *If $\mu \in p - BV$, then it follows that*

$$\operatorname{esssup}_{\gamma} \|\mu|_{\gamma}\|_{W_1} \leq A^{p-1} \|\mu\|_{p-BV}.$$

See [4], Lemma 2 for a proof in the case of real functions which also works in our case.

We now prove a sort of Helly selection principle for sequences of *positive* measures with bounded variation. This principle will be used, together with a regularization inequality, proved in next section, to get information on the variation of invariant measures. First we need a preliminary lemma:

Lemma 15 *If μ_n is a sequence of positive measures such that for each n , $\|\mu_n\|_{1^n} \leq C$, $var_p(\mu_n) \leq M$, and $\mu_n|_{\gamma} \rightarrow \mu|_{\gamma}$ for a.e. γ , in the Wasserstein distance, then*

$$\|\mu\|_{1^n} \leq C, var_p(\mu) \leq M.$$

Proof. Let us consider the $\|\cdot\|_1$ norm. Since by Remark 14 it holds $\|\mu_n|_\gamma\|_W \leq A^{p-1}(C+M) \forall \gamma$, by the dominated convergence theorem, $\|\mu\|_1 \leq C$. Let us now consider the oscillation. We have that $\liminf_{n \rightarrow \infty} \text{osc}(\mu_n, x, r) \geq \text{osc}(\mu, x, r)$ for all x, r . Indeed, consider a small $\epsilon \geq 0$. Because of the definition of $\text{osc}()$, we have that for all $l \geq \text{osc}(\mu, x, r) - \epsilon$, there are positive measure sets A_1 and A_2 such that $W_1(\mu|_{\gamma_1}, \mu|_{\gamma_2}) \geq l$ for all $\gamma_1 \in A_1, \gamma_2 \in A_2$. Since $\mu_n|_\gamma \rightarrow \mu|_\gamma$ for a.e. γ , if n is big enough there must be sets A_1^n and A_2^n of positive measure, such that $W_1(\mu_n|_{\gamma_1}, \mu_n|_{\gamma_2}) \geq l$ for all $\gamma_1 \in A_1^n, \gamma_2 \in A_2^n$, by this $\text{osc}(\mu_n, x, r) \geq l$, then for all x, r , $\liminf_{n \rightarrow \infty} \text{osc}(\mu_n, x, r) \geq \text{osc}(\mu, x, r)$. By Fatou's Lemma, $\liminf_{n \rightarrow \infty} \int_I r^{-p} \text{osc}(\mu_n, x, r) dx \geq \int_I r^{-p} \text{osc}(\mu, x, r) dx = \text{var}_p(\mu, r)$. From which the statement follows directly. ■

Theorem 16 (Helly-selection-like theorem) *Let μ_n be a sequence of probability measures on X such that $\|\mu_n\|_{p-BV} \leq M$ for some $M \geq 0$. Then there is μ with $\|\mu\|_{p-BV} \leq M$ and subsequence μ_{n_k} such that*

$$\|\mu_{n_k} - \mu\|_1 \rightarrow 0.$$

Proof. Let us discretize in the vertical direction. Let us consider a continuous projection of the space of probability measures on M on a finite dimensional space $\pi_{y,\delta} : PM(M) \rightarrow U_\delta$ (U_δ is finite dimensional). Suppose π_δ is such that $\|\pi_{y,\delta}(\nu) - \nu\|_{W_1} \leq C\delta, \forall \nu \in PM(M)$ (such projection can be constructed discretizing the space by a partition of unity made of Lipschitz functions with support on sets whose diameter is smaller than δ , see [23] for example). Let us consider the natural extension of this projection to the whole $\mathcal{L}^1(X)$ space $\pi_\delta : \mathcal{L}^1(X) \rightarrow \mathcal{L}^1(X)$, defined by $\pi_\delta(\mu)|_\gamma = \pi_{y,\delta}(\mu|_\gamma)$.

Let us consider the sequence $\pi_\delta(\mu_n)$. We have $\|\pi_\delta(\mu_n)\|_{p-BV} \leq K_\delta M$ where K_δ is the modulus of continuity of π_δ . Indeed

$$\text{esssup}_{\gamma_2, \gamma_1 \in B(x,r)} (W_1(\pi_\delta(\mu)|_{\gamma_1}, \pi_\delta(\mu)|_{\gamma_2})) \leq K_\delta \text{esssup}_{\gamma_2, \gamma_1 \in B(x,r)} (W_1(\mu|_{\gamma_1}, \mu|_{\gamma_2})).$$

Since after projecting we now are in a space of functions with values in a finite dimensional space, to the sequence $\pi_\delta(\mu_n)$ we can apply the classical Helly selection theorem and get that there is a limit measure μ_δ and a sub sequence n_k such that $\pi_\delta(\mu_{n_k}) \rightarrow \mu_\delta$ in \mathcal{L}^1 and $\pi_\delta(\mu_{n_k})|_\gamma \rightarrow \mu_\delta|_\gamma$ almost everywhere. Let us consider a sequence $\delta_i \rightarrow 0$ and select inductively at every step from the previous selected subsequence μ_l such that $\pi_{\delta_{i-1}}(\mu_l) \rightarrow \mu_{\delta_{i-1}}$ a further subsequence μ_{l_k} , such that $\pi_{\delta_i}(\mu_{l_k}) \rightarrow \mu_{\delta_i}$ in \mathcal{L}^1 and almost everywhere. Since $\forall \gamma$ and $m \leq i$, $\|\pi_{\delta_m}(\mu_{l_k}|_\gamma) - \mu_{l_k}|_\gamma\|_{W_1} \leq C\delta_m$, it holds that for different $\delta_m, \delta_j \geq \delta_i$

$$\|\pi_{\delta_i}(\mu_{l_k}|_\gamma) - \pi_{\delta_j}(\mu_{l_k}|_\gamma)\|_{W_1} \leq C(\delta_i + \delta_j)$$

and then $\forall \gamma$

$$\|\mu_{\delta_m}|_\gamma - \mu_{\delta_j}|_\gamma\|_{W_1} \leq C(\delta_m + \delta_j + \delta_i).$$

Since μ_{δ_n} are positive measures, this shows that there is a μ such that $\mu_{\delta_i} \rightarrow \mu$ in \mathcal{L}^1 and $\mu_{\delta_i}|_\gamma \rightarrow \mu|_\gamma$ almost everywhere, uniformly in γ . This shows

that a further subsequence μ_{n_j} can be selected in a way that $\pi_{\delta_i}(\mu_{n_j}) \xrightarrow{j \rightarrow \infty} \mu_{\delta_i}$ for all i , and $\mu_{n_j} \xrightarrow{j \rightarrow \infty} \mu$ in \mathcal{L}^1 and almost everywhere. Applying Lemma 15 we get $\|\mu\|_{p-BV} \leq M$. ■

4 A regularization inequality

In this section we prove an inequality, showing that iterates of a bounded variation positive measure are of uniform bounded variation. This will play the role of a Lasota Yorke inequality. A consequence will be a bound on the variation of invariant measures in \mathcal{L}^1 . This will be used when applying Theorem 3 to provide the estimate needed at Item 2 of. The following regularization inequality can be proved.

Proposition 17 *Let F be a skew product map satisfying assumptions Sk1, ..., Sk3 and let us suppose that μ is a positive measure. Let $p \leq \xi$ (the Hölder exponent as in Sk1). It holds*

$$\text{var}_p(L_F \mu) \leq \lambda^p \alpha \text{var}_p(\mu) + (\hat{H} \|\mu_x\|_\infty + 3q\alpha C_h A^{\xi-p} \|\mu_x\|_\infty).$$

We recall that here μ_x is the marginal of the disintegration of μ (see Equation 7) and $\|\mu_x\|_\infty$ is the supremum norm for its density.

Proof. By the Perron Frobenius like formula (Lemma 12)

$$(L_F \mu)|_\gamma = \sum_{i=1}^q \frac{F_{T_i^{-1}(\gamma)}^* \mu|_{T_i^{-1}(\gamma)}}{|T_i' \circ T_i^{-1}(\gamma)|} \text{ for almost all } \gamma \in [0, 1] \quad (13)$$

we have to estimate

$$\begin{aligned} I &: = \sup_{r \leq A} \frac{1}{r^p} \int \text{esssup}_{\gamma_2, \gamma_1 \in B(x, r)} \|(L_F \mu)|_{\gamma_1} - (L_F \mu)|_{\gamma_2}\|_{W_1} dm(x) \\ &= \sup_{r \leq A} \frac{1}{r^p} \int \text{esssup}_{\gamma_2, \gamma_1 \in B(x, r)} \left\| \sum_{i=1}^q \left(\frac{F_{T_i^{-1}(\gamma_1)}^* \mu|_{T_i^{-1}(\gamma_1)}}{|T_i' \circ T_i^{-1}(\gamma_1)|} - \frac{F_{T_i^{-1}(\gamma_2)}^* \mu|_{T_i^{-1}(\gamma_2)}}{|T_i' \circ T_i^{-1}(\gamma_2)|} \right) \right\|_{W_1} dm(x). \end{aligned}$$

By the triangular inequality

$$\begin{aligned} I &\leq \sup_{r \leq A} \sum_{i=1}^q \frac{1}{r^p} \int \text{esssup}_{\gamma_2, \gamma_1 \in B(x, r)} \left\| \frac{F_{T_i^{-1}(\gamma_1)}^* \mu|_{T_i^{-1}(\gamma_1)} - F_{T_i^{-1}(\gamma_2)}^* \mu|_{T_i^{-1}(\gamma_2)}}{|T_i' \circ T_i^{-1}(\gamma_1)|} \right\|_{W_1} dm \\ &\quad + \sup_{r \leq A} \sum_{i=1}^q \frac{1}{r^p} \int \text{esssup}_{\gamma_2, \gamma_1 \in B(x, r)} \left\| F_{T_i^{-1}(\gamma_2)}^* \mu|_{T_i^{-1}(\gamma_2)} \left(\frac{1}{|T_i' \circ T_i^{-1}(\gamma_1)|} - \frac{1}{|T_i' \circ T_i^{-1}(\gamma_2)|} \right) \right\|_{W_1} dm. \end{aligned}$$

Recalling that $\frac{1}{|T'_i \circ T_i^{-1}(\gamma_2)|} - \frac{1}{|T'_i \circ T_i^{-1}(\gamma_1)|} \leq C_h d(\gamma_1, \gamma_2)^\xi$, then

$$\begin{aligned} I &\leq \sup_{r \leq A} \sum_{i=1}^q \frac{1}{r^p} \int \left(\frac{1}{|T'_i \circ T_i^{-1}(x)|} + C_h r^\xi \right) \operatorname{esssup}_{\gamma_2, \gamma_1 \in B(x, r)} \|F_{T_i^{-1}(\gamma_1)}^* \mu|_{T_i^{-1}(\gamma_1)} - F_{T_i^{-1}(\gamma_2)}^* \mu|_{T_i^{-1}(\gamma_2)}\|_{W_1} dm \\ &\quad + \sup_{r \leq A} \sum_{i=1}^q \frac{1}{r^p} \int C_h r^\xi \operatorname{esssup}_{\gamma_2} \|F_{T_i^{-1}(\gamma_2)}^* \mu|_{T_i^{-1}(\gamma_2)}\|_{W_1} dm. \end{aligned}$$

And

$$\begin{aligned} I &\leq \sup_{r \leq A} \sum_{i=1}^q \frac{1}{r^p} \int \frac{1}{|T'_i \circ T_i^{-1}(x)|} \operatorname{esssup}_{\gamma_2, \gamma_1 \in B(x, r)} \|F_{T_i^{-1}(\gamma_1)}^* \mu|_{T_i^{-1}(\gamma_1)} - F_{T_i^{-1}(\gamma_2)}^* \mu|_{T_i^{-1}(\gamma_2)}\|_{W_1} dm \\ &\quad + \sup_{r \leq A} \sum_{i=1}^q \frac{1}{r^p} \int C_h r^\xi \operatorname{esssup}_{\gamma_2, \gamma_1 \in B(x, r)} \|F_{T_i^{-1}(\gamma_1)}^* \mu|_{T_i^{-1}(\gamma_1)} - F_{T_i^{-1}(\gamma_2)}^* \mu|_{T_i^{-1}(\gamma_2)}\|_{W_1} dm \\ &\quad + \sup_{r \leq A} \sum_{i=1}^q \frac{1}{r^p} \int C_h r^\xi \operatorname{esssup}_{\gamma_2} \|F_{T_i^{-1}(\gamma_2)}^* \mu|_{T_i^{-1}(\gamma_2)}\|_{W_1} dm. \end{aligned}$$

Hence

$$\begin{aligned} I &\leq \sup_{r \leq A} \sum_{i=1}^q \frac{1}{r^p} \int \frac{1}{|T'_i \circ T_i^{-1}(x)|} \operatorname{esssup}_{\gamma_2, \gamma_1 \in B(x, r)} \|F_{T_i^{-1}(\gamma_1)}^* \mu|_{T_i^{-1}(\gamma_1)} - F_{T_i^{-1}(\gamma_1)}^* \mu|_{T_i^{-1}(\gamma_2)}\|_{W_1} dm \\ &\quad + \sup_{r \leq A} \sum_{i=1}^q \frac{1}{r^p} \int \frac{1}{|T'_i \circ T_i^{-1}(x)|} \operatorname{esssup}_{\gamma_2, \gamma_1 \in B(x, r)} \|F_{T_i^{-1}(\gamma_1)}^* \mu|_{T_i^{-1}(\gamma_2)} - F_{T_i^{-1}(\gamma_2)}^* \mu|_{T_i^{-1}(\gamma_2)}\|_{W_1} dm \\ &\quad + 3 \sup_{r \leq A} \sum_{i=1}^q \frac{1}{r^p} \int C_h r^\xi \operatorname{esssup}_{\gamma_2} \|F_{T_i^{-1}(\gamma_2)}^* \mu|_{T_i^{-1}(\gamma_2)}\|_{W_1} dm \\ &= I_a + I_b + I_c \end{aligned}$$

Now

$$I_a \leq \sup_{r \leq A} \sum_{i=1}^q \frac{1}{r^p} \int \frac{1}{|T'_i \circ T_i^{-1}(x)|} \operatorname{esssup}_{\gamma_2, \gamma_1 \in B(x, r)} \|F_{T_i^{-1}(\gamma_1)}^* (\mu|_{T_i^{-1}(\gamma_1)} - \mu|_{T_i^{-1}(\gamma_2)})\|_{W_1} dx.$$

We recall that by Lemma 13 $\|F_\gamma^* \mu\|_{W_1} \leq \alpha \|\mu\|_{W_1}$ then

$$\begin{aligned} I_a &\leq \sup_{r \leq A} \sum_{i=1}^q \frac{1}{r^p} \int \frac{1}{|T'_i \circ T_i^{-1}(x)|} \operatorname{esssup}_{\gamma_2, \gamma_1 \in B(x, r)} \alpha \|\mu|_{T_i^{-1}(\gamma_1)} - \mu|_{T_i^{-1}(\gamma_2)}\|_{W_1} dx \\ &\leq \sup_{r \leq A} \sum_{i=1}^q \frac{1}{r^p} \int_{I_i} \operatorname{esssup}_{y_1, y_2 \in B(x, \lambda r)} \alpha \|\mu|_{y_1} - \mu|_{y_2}\|_{W_1} dx \\ &= \lambda^p \sup_{h \leq \lambda A} \sum_{i=1}^q \frac{1}{h^p} \int_{I_i} \operatorname{esssup}_{y_1, y_2 \in B(x, h)} \alpha \|\mu|_{y_1} - \mu|_{y_2}\|_{W_1} dx \end{aligned}$$

and

$$\text{var}_p(L_F \mu) \leq \lambda^p \alpha \text{var}_p(\mu) + (I_b + I_c).$$

By Sk3

$$\begin{aligned} I_b &= \sup_{r \leq A} \sum_i \frac{1}{r^p} \int \frac{1}{|T'_i \circ T_i^{-1}(x)|} \text{esssup}_{y_1, y_2 \in B(x, r)} \|F_{T_i^{-1}(\gamma_1)}^* \mu|_{T_i^{-1}(\gamma_2)} - F_{T_i^{-1}(\gamma_2)}^* \mu|_{T_i^{-1}(\gamma_2)}\|_{W_1} dm \\ &\leq \hat{H} \|\mu_x\|_\infty. \end{aligned}$$

Now, let us remark that since we are working with positive measures

$$\text{esssup}_{\gamma_2} \|F_{T_i^{-1}(\gamma_2)}^* \mu|_{T_i^{-1}(\gamma_2)}\|_{W_1} \leq \alpha \|\mu_x\|_\infty,$$

then

$$\begin{aligned} I_c &= 3 \sup_{r \leq A} \sum_{i=1}^q \frac{1}{r^p} \int C_h r^\xi \text{esssup}_{\gamma_2} \|F_{T_i^{-1}(\gamma_2)}^* \mu|_{T_i^{-1}(\gamma_2)}\|_{W_1} dm \\ &\leq 3 \sup_{r \leq A} \sum_{i=1}^q \frac{1}{r^p} \int C_h r^\xi \alpha \|\mu_x\|_\infty dm \end{aligned} \quad (14)$$

$$\leq 3q C_h A^{\xi-p} \alpha \|\mu_x\|_\infty \quad (15)$$

Summarizing

$$\text{var}_p(L_F \mu) \leq \lambda^p \alpha \text{var}_p(\mu) + (\hat{H} \|\mu_x\|_\infty + 3q \alpha C_h A^{\xi-p} \|\mu_x\|_\infty).$$

■

Remark 18 By Equation 5 it holds that for each n , $\|L^n \mu_x\|_\infty \leq C_\mu := A^{p-1}(A_T \lambda \|\mu_x\|_{BV} + B_T \|\mu_x\|_1 + 1)$. Iterating the last inequality we obtain

$$\begin{aligned} \text{var}_p(L_F^n \mu) &\leq (\lambda^p \alpha) \text{var}_p(L^{n-1} \mu) + (\hat{H} + 3q \alpha C_h A^{\xi-p}) C_\mu \\ &\leq \dots \\ &\leq (\lambda^p \alpha)^n \text{var}_p(\mu) + \frac{\hat{H} + 3q \alpha C_h A^{\xi-p}}{1 - \lambda^p \alpha} C_\mu \end{aligned} \quad (16)$$

By the Helly-like selection principle (Theorem 16) we then have

Proposition 19 In a system as above there is at least an invariant positive measure in $p - \mathcal{BV}$. For every such invariant measure μ

$$\text{var}_p(\mu) \leq \frac{B_T(\hat{H} + 3q \alpha C_h A^{\xi-p})}{1 - \lambda^p \alpha} \|\mu\|_1^n.$$

Proof. We consider the sequence of positive measures $\mu_n = \frac{1}{n} \sum L_F^n m$. By Equation 16, this sequence has uniformly bounded variation. Applying Theorem

16, we deduce the existence of an invariant measure μ in $p - \mathcal{BV}$. By the Lasota Yorke inequality relative to the map T , we have that

$$A^{p-1}B_T\|\mu\|_1 \geq A^{p-1}B_T\|\mu_x\|_1 \geq A^{p-1}\|\mu_x\|_{BV} \geq \|\mu_x\|_\infty.$$

This gives that

$$\text{var}_p(\mu) = \text{var}_p(L_F\mu) \leq \lambda^p \alpha \text{var}_p(\mu) + B_T\|\mu\|_1 A^{p-1}(\hat{H} + 3\alpha q C_h A^{\xi-p})$$

from which we get the statement. ■

5 Distance between the operators and a general statement for skew products

Here we consider a suitable class of perturbations of a map satisfying $Sk1, \dots, Sk3$ such that the associated transfer operators are near in the strong-weak topology, providing one of the estimate needed to apply Theorem 3 (item 4). In this section and in the following we set $p = 1$. Let us define Skorokhod distance between two piecewise expanding maps T_1 and T_2 as:

$$d_S(T_1, T_2) = \inf \left\{ \begin{array}{l} \epsilon > 0 : \exists A_1 \subseteq I \text{ and } \exists \sigma : I \rightarrow I \text{ such that } m(A_1) \geq 1 - \epsilon, \\ \sigma \text{ is a diffeomorphism, } T_1|_{A_1} = T_2 \circ \sigma|_{A_1} \\ \text{and } \forall x \in A_1, |\sigma(x) - x| \leq \epsilon, \left| \frac{1}{\sigma'(x)} - 1 \right| \leq \epsilon \end{array} \right\}. \quad (17)$$

It holds that one dimensional piecewise expanding maps which are near in the Skorokhod distance also have transfer operators which are near as operators from BV to L^1 (see [12], Lemma 11.2.1): there is $C_{Sk} \geq 0$ such that for each pair of piecewise expanding maps T_1, T_2

$$\|L_{T_0} - L_{T_\delta}\|_{BV \rightarrow L^1} \leq C_{Sk} d_S(T_1, T_2). \quad (18)$$

Let us see a statement of this kind for our skew products.

Proposition 20 *Let $F_\delta = (T_\delta, G_\delta)$, $0 \leq \delta \leq D$ be a family of maps satisfying $Sk1, \dots, Sk3$ uniformly with $\xi = 1$ and:*

1. *For each $\delta \leq D$, $d_S(T_0, T_\delta) \leq \delta$. (thus for each δ there is a set A_1 as in the definition of the Skorokhod distance)*
2. *For each $\delta \leq D$ there is a set A_2 such that $m(A_2) \geq 1 - \delta$ and for all $x \in A_2, y \in M : |G_0(x, y) - G_\delta(x, y)| \leq \delta$.*

Let us denote by F_δ^ the transfer operators of F_δ and by f_δ a family of probability measures with uniformly bounded variation*

$$\text{var}_1(f_\delta) \leq M_2.$$

Then, there is a constants C_1 such that for δ small enough

$$\|(F_0^* - F_\delta^*)f_\delta\|_1 \leq C_1 \delta (M_2 + 1).$$

Proof. Let us set $A = A_1 \cup A_2$. Note that $m(A^c) \leq 2\delta$. Let us estimate

$$\begin{aligned} \|(F_0^* - F_\delta^*)f_\delta\|_1 &= \int_I \|(F_0^* f_\delta - F_\delta^* f_\delta)|_\gamma\|_W dm(\gamma) \\ &= \int_I \|F_0^*(1_A f_\delta)|_\gamma - F_\delta^*(1_A f_\delta)|_\gamma\|_W dm(\gamma) + \int_I \|F_0^*(1_{A^c} f_\delta)|_\gamma - F_\delta^*(1_{A^c} f_\delta)|_\gamma\|_W dm(\gamma). \end{aligned} \quad (19)$$

By the assumptions, for a.e. γ , $\|f_\delta|_\gamma\|_W \leq (M_2 + 1)$ and $\|1_{A^c} f_\delta\|_1 \leq (M_2 + 1)\delta$. Since F^* is α -Lipschitz for the \mathcal{L}^1 norm then

$$\int_I \|F_0^*(1_{A^c} f_\delta)|_\gamma - F_\delta^*(1_{A^c} f_\delta)|_\gamma\|_W dm(\gamma) \leq 2\alpha(M_2 + 1)\delta.$$

Let us now estimate the first summand of 19. Let us set $\mu = 1_A f_\delta$ and let us estimate the integral

$$\int \|(F_0^* \mu - F_\delta^* \mu)|_\gamma\|_W dm(\gamma).$$

Let us denote by $T_{0,i}$, with $0 \leq i \leq q$ the branches of T_0 defined in the sets P_i , partition of I , and set $T_{\delta,i} = T_\delta|_{P_i \cap A}$ these functions will play the role of the branches for T_δ . Since in A , $T_0 = T_\delta \circ \sigma_\delta$ (where σ_δ is the diffeomorphism in the definition of the Skorokhod distance), then $T_{\delta,i}$ are invertible. Then

$$(F_0^* \mu - F_\delta^* \mu)|_\gamma = \sum_{i=1}^q \frac{F_{0,T_{0,i}^{-1}(\gamma)}^* \mu|_{T_{0,i}^{-1}(\gamma)} \chi_{T_0(P_i \cap A)}}{|T'_{0,i}(T_{0,i}^{-1}(\gamma))|} - \sum_{i=1}^q \frac{F_{\delta,T_{\delta,i}^{-1}(\gamma)}^* \mu|_{T_{\delta,i}^{-1}(\gamma)} \chi_{T_\delta(P_i \cap A)}}{|T'_{\delta,i}(T_{\delta,i}^{-1}(\gamma))|} \quad \mu_x - a.e. \gamma \in I,$$

Let us now consider $T_0(P_i \cap A)$ and $T_\delta(P_i \cap A)$, and remark that $T_0(P_i \cap A) = \sigma_\delta(T_\delta(P_i \cap A))$ and σ_δ is a diffeomorphism near to the identity. Let us denote $B_i = T_0(P_i \cap A) \cap T_\delta(P_i \cap A)$, $C_i = T_0(P_i \cap A) \triangle T_\delta(P_i \cap A)$.

$$\begin{aligned} \int_I \|(F_0^* \mu - F_\delta^* \mu)|_\gamma\|_W dm(\gamma) &\leq \int_I \left\| \sum_{i=1}^q \frac{F_{0,T_{0,i}^{-1}(\gamma)}^* \mu|_{T_{0,i}^{-1}(\gamma)} \chi_{B_i}}{|T'_{0,i}(T_{0,i}^{-1}(\gamma))|} - \sum_{i=1}^q \frac{F_{\delta,T_{\delta,i}^{-1}(\gamma)}^* \mu|_{T_{\delta,i}^{-1}(\gamma)} \chi_{B_i}}{|T'_{\delta,i}(T_{\delta,i}^{-1}(\gamma))|} \right\|_W dm \\ &\quad + \int_I \left\| \sum_{i=1}^q \frac{F_{0,T_{0,i}^{-1}(\gamma)}^* \mu|_{T_{0,i}^{-1}(\gamma)} \chi_{C_i}}{|T'_{0,i}(T_{0,i}^{-1}(\gamma))|} - \sum_{i=1}^q \frac{F_{\delta,T_{\delta,i}^{-1}(\gamma)}^* \mu|_{T_{\delta,i}^{-1}(\gamma)} \chi_{C_i}}{|T'_{\delta,i}(T_{\delta,i}^{-1}(\gamma))|} \right\|_W dm. \end{aligned} \quad (20)$$

And since there is K_1 such that $m(C_i) \leq K_1 \delta$, then

$$\int_I \left\| \sum_{i=1}^q \frac{F_{0,T_{0,i}^{-1}(\gamma)}^* \mu|_{T_{0,i}^{-1}(\gamma)} \chi_{C_i}}{|T'_{0,i}(T_{0,i}^{-1}(\gamma))|} - \sum_{i=1}^q \frac{F_{\delta,T_{\delta,i}^{-1}(\gamma)}^* \mu|_{T_{\delta,i}^{-1}(\gamma)} \chi_{C_i}}{|T'_{\delta,i}(T_{\delta,i}^{-1}(\gamma))|} \right\|_W dm \leq qK_1(M_2 + 1)\delta.$$

Now we have to consider the first summand of 20. We have

$$\begin{aligned}
& \int_I \left\| \sum_{i=1}^q \frac{F_{0,T_{0,i}^{-1}(\gamma)}^* \mu|_{T_{0,i}^{-1}(\gamma)} \chi_{B_i}}{|T'_{0,i}(T_{0,i}^{-1}(\gamma))|} - \sum_{i=1}^q \frac{F_{\delta,T_{\delta,i}^{-1}(\gamma)}^* \mu|_{T_{\delta,i}^{-1}(\gamma)} \chi_{B_i}}{|T'_{\delta,i}(T_{\delta,i}^{-1}(\gamma))|} \right\|_W dm \\
& \leq \int_I \left\| \sum_{i=1}^q \frac{F_{0,T_{0,i}^{-1}(\gamma)}^* \mu|_{T_{0,i}^{-1}(\gamma)} \chi_{B_i}}{|T'_{0,i}(T_{0,i}^{-1}(\gamma))|} - \sum_{i=1}^q \frac{F_{\delta,T_{\delta,i}^{-1}(\gamma)}^* \mu|_{T_{0,i}^{-1}(\gamma)} \chi_{B_i}}{|T'_{\delta,i}(T_{\delta,i}^{-1}(\gamma))|} \right\|_W dm \\
& \quad + \int_I \left\| \sum_{i=1}^q \frac{F_{\delta,T_{\delta,i}^{-1}(\gamma)}^* \mu|_{T_{0,i}^{-1}(\gamma)} \chi_{B_i}}{|T'_{\delta,i}(T_{\delta,i}^{-1}(\gamma))|} - \sum_{i=1}^q \frac{F_{\delta,T_{\delta,i}^{-1}(\gamma)}^* \mu|_{T_{\delta,i}^{-1}(\gamma)} \chi_{B_i}}{|T'_{\delta,i}(T_{\delta,i}^{-1}(\gamma))|} \right\|_W dm \\
& = \int_I I(\gamma) dm(\gamma) + \int_I II(\gamma) dm(\gamma).
\end{aligned}$$

The two summands will be treated separately.

$$\begin{aligned}
I(\gamma) &= \left\| \sum_{i=1}^q \frac{F_{0,T_{0,i}^{-1}(\gamma)}^* \mu|_{T_{0,i}^{-1}(\gamma)} \chi_{B_i}}{|T'_{0,i}(T_{0,i}^{-1}(\gamma))|} - \sum_{i=1}^q \frac{F_{\delta,T_{\delta,i}^{-1}(\gamma)}^* \mu|_{T_{0,i}^{-1}(\gamma)} \chi_{B_i}}{|T'_{\delta,i}(T_{\delta,i}^{-1}(\gamma))|} \right\|_W \\
&\leq \left\| \sum_{i=1}^q \frac{F_{0,T_{0,i}^{-1}(\gamma)}^* \mu|_{T_{0,i}^{-1}(\gamma)} \chi_{B_i}}{|T'_{0,i}(T_{0,i}^{-1}(\gamma))|} - \sum_{i=1}^q \frac{F_{\delta,T_{\delta,i}^{-1}(\gamma)}^* \mu|_{T_{0,i}^{-1}(\gamma)} \chi_{B_i}}{|T'_{0,i}(T_{0,i}^{-1}(\gamma))|} \right\|_W \\
&\quad + \left\| \sum_{i=1}^q \frac{F_{\delta,T_{\delta,i}^{-1}(\gamma)}^* \mu|_{T_{0,i}^{-1}(\gamma)} \chi_{B_i}}{|T'_{0,i}(T_{0,i}^{-1}(\gamma))|} - \sum_{i=1}^q \frac{F_{\delta,T_{\delta,i}^{-1}(\gamma)}^* \mu|_{T_{\delta,i}^{-1}(\gamma)} \chi_{B_i}}{|T'_{\delta,i}(T_{\delta,i}^{-1}(\gamma))|} \right\|_W \\
&= I_a(\gamma) + I_b(\gamma).
\end{aligned}$$

Since f_δ is a probability measure it holds posing $\beta = T_{0,i}^{-1}(\gamma)$

$$\begin{aligned}
\int I_a(\gamma) dm &= \int \left\| \sum_{i=1}^q \frac{F_{0,T_{0,i}^{-1}(\gamma)}^* \mu|_{T_{0,i}^{-1}(\gamma)} \chi_{B_i}}{|T'_{0,i}(T_{0,i}^{-1}(\gamma))|} - \sum_{i=1}^q \frac{F_{\delta,T_{\delta,i}^{-1}(\gamma)}^* \mu|_{T_{0,i}^{-1}(\gamma)} \chi_{B_i}}{|T'_{0,i}(T_{0,i}^{-1}(\gamma))|} \right\|_W dm(\gamma) \\
&\leq \int \sum_{i=1}^q \left\| \frac{F_{0,T_{0,i}^{-1}(\gamma)}^* \mu|_{T_{0,i}^{-1}(\gamma)} \chi_{B_i}}{|T'_{0,i}(T_{0,i}^{-1}(\gamma))|} - \frac{F_{\delta,T_{\delta,i}^{-1}(\gamma)}^* \mu|_{T_{0,i}^{-1}(\gamma)} \chi_{B_i}}{|T'_{0,i}(T_{0,i}^{-1}(\gamma))|} \right\|_W dm \\
&\leq \sum_{i=1}^q \int \left\| \frac{F_{0,T_{0,i}^{-1}(\gamma)}^* \mu|_{T_{0,i}^{-1}(\gamma)} \chi_{B_i}}{|T'_{0,i}(T_{0,i}^{-1}(\gamma))|} - \frac{F_{\delta,T_{\delta,i}^{-1}(\gamma)}^* \mu|_{T_{0,i}^{-1}(\gamma)} \chi_{B_i}}{|T'_{0,i}(T_{0,i}^{-1}(\gamma))|} \right\|_W dm \\
&\leq \sum_{i=1}^q \int_{T_{0,i}^{-1}(B_i)} \left\| F_{0,\beta}^* \mu|_\beta - F_{\delta,T_{\delta,i}^{-1}(T_{0,i}(\beta))}^* \mu|_\beta \right\|_W dm(\beta)
\end{aligned}$$

Remark that $T_{0,i}^{-1}(B_i) \subseteq P_i \cap A$ and $T_{\delta,i}^{-1}(T_{0,i}(T_{0,i}^{-1}(B_i))) \subseteq P_i \cap A$. Since $|T_{\delta,i}(\beta) - T_{0,i}(\beta)| \leq \delta$ and $T_{0,i}^{-1}$ is a contraction, then $|T_{0,i}^{-1} \circ T_{\delta,i}(\beta) - \beta| \leq \delta$.

Then

$$\begin{aligned} \left\| F_{0,\beta}^* \mu|_\beta - F_{\delta,T_{\delta,i}^{-1}(T_{0,i}(\beta))}^* \mu|_\beta \right\|_W &\leq \left\| F_{0,\beta}^* \mu|_\beta - F_{\delta,\beta}^* \mu|_\beta \right\|_W \\ &\quad + \left\| F_{\delta,\beta}^* \mu|_\beta - F_{\delta,T_{\delta,i}^{-1}(T_{0,i}(\beta))}^* \mu|_\beta \right\|_W. \end{aligned}$$

By assumption (2),

$$\left\| F_{0,\beta}^* \mu|_\beta - F_{\delta,\beta}^* \mu|_\beta \right\|_W \leq \delta(M_2 + 1).$$

By assumption *Sk3*

$$\left\| F_{\delta,\beta}^* \mu|_\beta - F_{\delta,T_{\delta,i}^{-1}(T_{0,i}(\beta))}^* \mu|_\beta \right\|_W \leq \sup_{y \in M, x_1, x_2 \in B(\beta, \delta)} |G(x_1, y) - G(x_2, y)| (M_2 + 1).$$

Thus,

$$I_a(\gamma) \leq \delta^p(\hat{H} + 1)(M_2 + 1) + \delta(M_2 + 1).$$

To estimate $I_b(\gamma)$ we have:

$$\begin{aligned} I_b(\gamma) &= \left\| \sum_{i=1}^q \frac{F_{\delta,T_{\delta,i}^{-1}(\gamma)}^* \mu|_{T_{0,i}^{-1}(\gamma)} \chi_{B_i}}{|T'_{0,i}(T_{0,i}^{-1}(\gamma))|} - \sum_{i=1}^q \frac{F_{\delta,T_{\delta,i}^{-1}(\gamma)}^* \mu|_{T_{0,i}^{-1}(\gamma)} \chi_{B_i}}{|T'_{\delta,i}(T_{\delta,i}^{-1}(\gamma))|} \right\|_W \\ &\leq \sum_{i=1}^q \left| \frac{\chi_{B_i}(\gamma)}{|T'_{0,i}(T_{0,i}^{-1}(\gamma))|} - \frac{\chi_{B_i}(\gamma)}{|T'_{\delta,i}(T_{\delta,i}^{-1}(\gamma))|} \right| \left\| F_{\delta,T_{\delta,i}^{-1}(\gamma)}^* \mu|_{T_{0,i}^{-1}(\gamma)} \right\|_W \end{aligned}$$

and

$$\int I_b \, dm \leq |(P_{T_0} - P_{T_\delta})(1)| \alpha(M_2 + 1) + qK_1\delta.$$

by Equation 18 then

$$\int_{A_1} I_b(\gamma) \, dm(\gamma) \leq [C_{Sk} \alpha(M_2 + 1) + qK_1] \delta.$$

Now, let us estimate the integral of the second summand

$$II(\gamma) = \left\| \sum_{i=1}^q \frac{F_{\delta,T_{\delta,i}^{-1}(\gamma)}^* \mu|_{T_{0,i}^{-1}(\gamma)} \chi_{B_i}}{|T'_{\delta,i}(T_{\delta,i}^{-1}(\gamma))|} - \sum_{i=1}^q \frac{F_{\delta,T_{\delta,i}^{-1}(\gamma)}^* \mu|_{T_{\delta,i}^{-1}(\gamma)} \chi_{B_i}}{|T'_{\delta,i}(T_{\delta,i}^{-1}(\gamma))|} \right\|_W.$$

Then,

$$\begin{aligned} \int_I II(\gamma) \, dm(\gamma) &= \int_I \left\| \sum_{i=1}^q \frac{F_{\delta,T_{\delta,i}^{-1}(\gamma)}^* \mu|_{T_{0,i}^{-1}(\gamma)} \chi_{B_i}}{|T'_{\delta,i}(T_{\delta,i}^{-1}(\gamma))|} - \sum_{i=1}^q \frac{F_{\delta,T_{\delta,i}^{-1}(\gamma)}^* \mu|_{T_{\delta,i}^{-1}(\gamma)} \chi_{B_i}}{|T'_{\delta,i}(T_{\delta,i}^{-1}(\gamma))|} \right\|_W dm(\gamma) \\ &\leq \sum_{i=1}^q \int_{B_i} \frac{1}{|T'_{\delta,i}(T_{\delta,i}^{-1}(\gamma))|} \left\| F_{\delta,T_{\delta,i}^{-1}(\gamma)}^* (\mu|_{T_{0,i}^{-1}(\gamma)} - \mu|_{T_{\delta,i}^{-1}(\gamma)}) \right\|_W dm(\gamma) \\ &\leq \sum_{i=1}^q \int_{B_i} \frac{\alpha}{|T'_{\delta,i}(T_{\delta,i}^{-1}(\gamma))|} \left\| \mu|_{T_{0,i}^{-1}(\gamma)} - \mu|_{T_{\delta,i}^{-1}(\gamma)} \right\|_W dm(\gamma) \end{aligned}$$

Let us consider the change of variable $\gamma = T_{\delta,i}(\beta)$, then

$$\int_I II(\gamma) dm(\gamma) \leq \alpha \sum_{i=1}^q \int_{T_{\delta,i}^{-1}(B_i)} \left\| |\mu|_{T_{\delta,i}^{-1}(T_{\delta,i}(\beta))} - \mu|_{\beta} \right\|_W dm(\beta).$$

Since $|T_{\delta,i}(\beta) - T_{0,i}(\beta)| \leq \delta$ and $T_{0,i}^{-1}$ is a contraction, then $|T_{0,i}^{-1} \circ T_{\delta,i}(\beta) - \beta| \leq \delta$ then

$$\int_I II(\gamma) dm(\gamma) \leq \alpha \int \sup_{x,y \in B(\beta,\delta)} (||\mu|_x - \mu|_y||_W) dm(\beta) \leq \alpha \int osc(\mu, \beta, \delta) d\beta$$

and then

$$\int_I II(\gamma) dm(\gamma) \leq \alpha 2\delta(M_2 + 1).$$

Summing all, the statement is proved. ■

The last statement, together with the results of the previous sections allows to prove the following quantitative statement for skew product maps.

Proposition 21 *Consider a family of skew product maps $F_\delta = (T_\delta, G_\delta)$, $0 \leq \delta \leq D$ satisfying Sk1, ..., Sk3 uniformly, with $\xi = 1$, and let $f_\delta \in \mathcal{L}^1$ invariant probability measures of F_δ , suppose:*

1. *There is $\phi \in C^0(\mathbb{R})$, $\phi(t)$ decreasing to 0 as $t \rightarrow \infty$ such that L_{F_0} has convergence to equilibrium with respect to norms $|| \cdot ||_{1-BV}$, $|| \cdot ||_{1'}$ and speed ϕ ;*
2. *there is $\tilde{C} \geq 0$ such that for each n*

$$||L_{F_0}^n||_{B_w \rightarrow B_w} \leq \tilde{C};$$

3. *for each $\delta \leq D$, $d_S(T_0, T_\delta) \leq \delta$;*
4. *for each $\delta \leq D$ there is a set A_2 such that $m(A_2) \geq 1 - \delta$ and for all $x \in A_2, y \in M : |G_0(x, y) - G_\delta(x, y)| \leq \delta$.*

Let $B = \frac{B_T(\hat{H} + 3q\alpha C_h)}{1 - \lambda^p \alpha} + 1$. Consider the function ψ defined as $\psi(x) = \frac{\phi(x)}{x}$, then

$$||f_\delta - f_0||_{1'} \leq 2\tilde{C}B^2C_1\delta(\psi^{-1}(\frac{\tilde{C}BC_1\delta}{2}) + 1).$$

where C_1 is the constant in the statement of Proposition 20.

Proof. The proof is a direct application of the estimates given in the previous section into Theorem 3. The quantity \tilde{M} appearing at Item 2 of Theorem 3 is estimated by Proposition 19:

$$\tilde{M} \leq \frac{B_T(\hat{H} + 3q\alpha C_h)}{1 - \lambda^p \alpha}.$$

By Proposition 20 the distance between the operators appearing at Item 4 of Theorem 3 is bounded by $\epsilon \leq C_1 \delta (M_2 + 1)$ Where M_2 bounds the strong norm of f_δ . ■

We remark that the quantitative stability is proved here in the $\|\cdot\|_1$ topology. This topology is strong enough to control the behavior of observables which are discontinuous along the preserved central foliation, see [10] for other results on quantitative stability of the statistical properties of discontinuous observables and related applications.

In the following section we show a class of nontrivial partially hyperbolic skew products having power law convergence to equilibrium and will apply this statement to these examples.

6 Application to slowly mixing toral extensions

To give an example of application of Proposition 21 to a class of nontrivial system, we consider a class of "partially hyperbolic" skew products with some discontinuities, having slow (power law) decay of correlations and convergence to equilibrium.

We will consider a class of skew products $F = (T, G)$ (piecewise constant toral extensions) defined as follows:

Te1 let $l \in \mathbb{N}$. We assume that T is the piecewise expanding map on $[0, 1]$ defined as

$$T(x) = lx \bmod(1);$$

Te2 the system is extended by a skew product to a system (X, F) where $X = [0, 1] \times \mathcal{T}^d$, where \mathcal{T}^d is the d dimensional torus and $F : X \rightarrow X$ is defined by

$$F(x, t) = (Tx, t + \theta \varphi(\omega)) \quad (21)$$

where $\theta = (\theta_1, \dots, \theta_d) \in \mathcal{T}^d$ and $\varphi = 1_I$ is the characteristic function of a set $I \subset [0, 1]$ which is an union of the sets P_i where the branches of T are defined. In this system the second coordinate is translated by θ if the first coordinate belongs to I .

We remark that on the system (X, F) the Lebesgue measure is invariant. We will suppose that θ is of finite Diophantine type. Let us recall the definition of Diophantine type for the linear approximation. The definition tests the possibility of approximating 0 by an integer linear combination of its components.

The notation $\|\cdot\|$ will indicate the distance to the nearest integer vector (or number) in \mathbb{R} , and $|k| = \sup_{0 \leq i \leq d} |k_i|$ indicates the supremum norm.

Definition 22 *The Diophantine type of $\theta = (\theta_1, \dots, \theta_d)$ for the linear approximation is*

$$\gamma_l(\theta) = \inf\{\gamma, \text{ s.t. } \exists c_0 > 0 \text{ s.t. } \|k \cdot \theta\| \geq c_0 |k|^{-\gamma} \forall 0 \neq k \in \mathbb{Z}^d\}.$$

6.1 The decay of correlations

In [31], it was observed that piecewise constant toral extensions cannot have exponential decay of correlations (in [30] by the way it is shown that for some piecewise constant $SU_2(\mathbb{C})$ extensions there can be exponential decay of correlations). Quantitative estimates for the speed of decay of correlations by the arithmetical properties of the angles, have been given in [21].

In this section we recall those results and see that the systems defined above have at least polynomial decay of correlations, while for some choice of the angles the speed of decay is proved to be actually polynomial.

Definition 23 (Decay of correlations) *Let $\phi, \psi : X \rightarrow \mathbb{R}$ be observables on X belonging to the Banach spaces B, B' , let ν be an invariant measure for T . Let $\Phi : \mathbb{N} \rightarrow \mathbb{R}$ such that $\Phi(n) \xrightarrow{n \rightarrow \infty} 0$. A system (X, T, ν) is said to have decay of correlations with speed Φ with respect to observables in B and B' if*

$$\left| \int \phi \circ T^n \psi d\nu - \int \phi d\nu \int \psi d\nu \right| \leq \|\phi\|_B \|\psi\|_{B'} \Phi(n) \quad (22)$$

where $\|\cdot\|_B, \|\cdot\|_{B'}$ are the norms in B and B' .

The decay of correlations depends on the class of observables considered. On the skew products satisfying conditions *Te1* and *Te2* as above, it is possible to establish an explicit upper bound for the rate of decay of correlations which depend on the linear type of the translation angle (see [21], Lemma11).

Proposition 24 *In the piecewise constant toral extensions described above, for Lipschitz observables the rate of decay is*

$$\Phi(n) = O(n^{-\frac{1}{2\gamma}})$$

for any $\gamma > \gamma_l(\theta)$.

For C^p, C^q observables, the rate of decay is

$$\Phi(n) = O(n^{-\frac{1}{2\gamma} \max(p, q, p+q-d)})$$

for any $\gamma > \gamma_l(\theta)$.

Remark 25 *We remark that the rate is actually polynomial in some cases. In [21], Section 5 (using a result of [24]) it is proved that if the Diophantine type is large, then the mixing rate of the systems we consider is actually slow, with a power law speed which depends on the Diophantine type. In a system satisfying (22), let the exponent of power law decay be defined by*

$$p = \liminf_{n \rightarrow \infty} \frac{-\log \Phi(n)}{\log n}.$$

Let us consider the skew product of the doubling map and a circle rotation endowed with the Lebesgue (invariant) measure. For this example the exponent p satisfies

$$\frac{1}{2\gamma(\theta)} \leq p \leq \frac{6}{\max(2, \gamma(\theta)) - 2}.$$

6.2 Convergence to equilibrium

We will use the decay of correlations of the toral extensions to get a convergence to equilibrium result with respect to the strong and weak norm of our anisotropic spaces. We have from Proposition 24 that for Lipschitz observables the rate of decay is $O(n^{-\frac{1}{2\gamma}})$ for any $\gamma > \gamma_t(\alpha)$ and for any Lipschitz observables with $\int f = 0$:

$$|\int g \circ F^n f \, dm| \leq C \|f\|_{lip} \|g\|_{lip} n^{-\frac{1}{2\gamma}}.$$

From this we will prove that

$$\|L^n \mu\|_{1^*} \leq C_4 n^{-\frac{1}{8\gamma}} \|\mu\|_{1-BV}.$$

For this purpose our strategy is to approximate a $1-BV$ measure μ which is meant to be iterated with a Lipschitz density and use the decay of correlations with Lipschitz observables to estimate its convergence to equilibrium. We remark that a statement of this kind extend greatly the kinds of measures which are meant to be iterated, as the space of $1-BV$ measures contains measures with singular behavior in the neutral direction.

The first step in the strategy is approximating the disintegration of μ with a kind of "piecewise constant one" in next Lemma.

Lemma 26 *Let us consider a uniform grid of size $\epsilon = \frac{1}{m}$, $m \in \mathbb{N}$, on the interval $[0, 1]$. Given a measure μ with $\|\mu\|_{1-BV} < \infty$. There is a measure μ_ϵ such that μ_ϵ is piecewise constant on the ϵ -grid ($\mu_\epsilon|_x$ is constant on each element of the grid as x varies) and*

$$var_1(\mu_\epsilon) \leq 2var_1(\mu), \quad \|\mu_\epsilon\|_1 \leq \|\mu\|_1$$

and

$$\|\mu - \mu_\epsilon\|_1 \leq 2\epsilon var_1(\mu).$$

Proof. Let us consider μ_ϵ defined averaging in the following way: let $x \in [0, 1]$ and I_i be the element of the ϵ -grid containing x . Then for a measurable set $A \subseteq M$ $\mu|_x(A)$ is defined as

$$\mu_\epsilon|_x(A) = \int_{I_i} \mu_\epsilon|_\gamma(A) d\gamma.$$

We remark that $\mu|_x - \mu_\epsilon|_x \leq osc(\epsilon, x_i(x), \mu)$ where $x_i(x)$ is the grid center closest to x , and $osc(\epsilon, x_i(x), \mu) \leq osc(2\epsilon, x, \mu)$ then

$$\begin{aligned} \int \|\mu|_x - \mu_\epsilon|_x\|_W &\leq 2\epsilon \frac{\int osc(2\epsilon, x, \mu)}{2\epsilon} \\ &\leq 2\epsilon \sup_{2\epsilon \leq A} \left(\frac{\int osc(2\epsilon, x, \mu)}{2\epsilon} \right) \\ &\leq 2\epsilon var_p(\mu) \end{aligned}$$

The other inequalities are analogous. ■

Proposition 27 *The convergence to equilibrium of a system satisfying Te1, Te2 can be estimated as*

$$\|L^n \nu\|_1 \leq C_4 n^{-\frac{1}{8\gamma}} \|\nu\|_{1-BV}.$$

Proof. Consider a $1-BV$ measure ν , without loss of generality we can suppose $\|\nu\|_{1-BV} = 1$. Let us approximate ν it with a Lipschitz measure. First let us approximate it with a piecewise constant measure ν_ϵ as before. We have

$$\|\nu - \nu_\epsilon\|_1 \leq 2\epsilon \text{var}_1(\nu) \leq 2\epsilon.$$

Let ν_i be such that $\nu_i = \nu_\epsilon|_{x_i}$ with x_i center of the ϵ grid as before, and f_i be the convolution $\gamma * \nu_i$ where γ is a ϵ_2^{-1} Lipschitz mollifier supported in $[-\epsilon_2, \epsilon_2]^d$. f_i is a ϵ_2^{-1} Lipschitz function. Let

$$f(x, y) = \begin{cases} f_i(y) & \text{if } |x - x_i| \leq (1 - \epsilon_2)\epsilon \\ \phi_i(x)f_i(y) + (1 - \phi_i(x))f_{i+1}(y) & \text{if } x_i + (1 - \epsilon_2)\epsilon \leq x \leq x_{i+1} - (1 - \epsilon_2)\epsilon \end{cases}$$

where ϕ_i is a linear function s.t. $\phi_i(x_i + (1 - \epsilon_2)\epsilon) = 0$ and $\phi_i(x_{i+1} - (1 - \epsilon_2)\epsilon) = 1$. We remark that f is $\sqrt{2}\epsilon_2^{-1}\epsilon^{-1}$ Lipschitz, $\int f \, dm = 0$. and $\|\nu_\epsilon - fm\|_{1^*} \leq 3\epsilon_2$. Hence

$$\|\nu - fm\|_{1^*} \leq 2\epsilon \text{var}_1(\nu) + 3\epsilon_2. \quad (23)$$

Since the convolution with a Lipschitz kernel is a weak contraction in the Wasserstein norm, applying Lemma 26 we get $\text{var}_1(fm) \leq 2\text{var}_1(\nu)$ and $\|f\|_{1^*} \leq \|\nu\|_{1^*}$. Now we apply Proposition 24 in an efficient way. The proposition concerns the behavior of the correlation of two observables. We will consider f as one of them, and the other will be constructed in a suitable way to get the desired statement.

Let f be the Lipschitz density found above. Let $\mu = L^n fm$. Let μ_ϵ its approximation defined as in Lemma 26 and $\mu_i = \mu_\epsilon|_{x_i}$. Consider 1-Lipschitz functions $l_i : \mathcal{T}^d \rightarrow \mathbb{R}$ such that $|\int l_i \mu_i - \|\mu_i\|_W| \leq \xi$, consider functions $h_i : [0, 1] \rightarrow \mathbb{R}$ such that $h_i = 1$ on the central third of the i interval of the ϵ -grid and zero elsewhere, and $\text{lip}(h_i) = 3\epsilon^{-1}$. Consider $g_i : X \rightarrow \mathbb{R}$ defined by $g_i(x, y) = l_i(y)h_i(x)$ and $g = \sum_i g_i$. By what is said above

$$\|\mu_\epsilon\|_1 \leq 3(\xi + \int g \mu_\epsilon)$$

and by Lemma 26, $\|L^n fm - \mu_\epsilon\|_1 \leq 2\epsilon \text{var}_1(\mu)$. Then

$$\begin{aligned} \|L^n(fm)\|_1 &\leq \|\mu_\epsilon\|_1 + 2\epsilon \text{var}_1(\mu) \\ &\leq 3(\xi + \int g \mu_\epsilon) + 2\epsilon \text{var}_1(\mu). \end{aligned}$$

Now consider $\int g L^n fm$. Since g is 1-Lipschitz in the y direction, we have that

$$|\int g L^n f - \int g \mu_\epsilon| \leq \|L^n fm - \mu_\epsilon\|_1 \leq 2\epsilon \text{var}_1(\mu)$$

and

$$\|L^n f\|_1 \leq 3(\xi + \int g dL^n f \mu_0 + 2\epsilon \text{var}_1(\mu)) + 2\epsilon \text{var}_1(\mu).$$

Now, since $\int f dm = 0$, by Proposition 24

$$|\int g dL^n(fm)| \leq C\|f\|_{lip}\|g\|_{lip}n^{-\frac{1}{2\gamma}}$$

then

$$\|L^n f\|_1 \leq 3(\xi + C\|f\|_{lip}3\epsilon^{-1}n^{-\frac{1}{2\gamma}} + 2\epsilon \text{var}_1(L^n(fm))) + 2\epsilon \text{var}_1(L^n(fm)).$$

We recall that the Lebesgue measure is invariant for the system. Then if K is a constant density such that $f+K \geq 0$ it holds $L^n(fm+Km) = L^n(fm)+Km$. It holds $\text{var}_1(L^n f) = \text{var}_1(L^n(f+K))$, since it is a positive measure, to $(f+K)m$ we can apply the regularization inequality. Setting $B = \frac{B_T(\lambda^p + \hat{H} + 3qC_h)}{1 - \lambda^p \alpha}$ we get

$$\text{var}_1(L^n(fm)) \leq \lambda^n \alpha^n \text{var}_1(f) + B(\|f\|_{1^n} + K) \quad (24)$$

since f is $\sqrt{2}\epsilon_2^{-1}\epsilon^{-1}$ -Lipschitz and $\|f\|_{1-BV}$ density, $\|f\|_\infty \leq \sqrt{2}\epsilon_2^{-1}\epsilon^{-1} + 1$, then $\text{var}(L^n(fm)) \leq \lambda^n \alpha^n \text{var}(f) + B(1 + \sqrt{2}\epsilon_2^{-1}\epsilon^{-1} + 1)$ and

$$\begin{aligned} \|L^n(fm)\|_1 &\leq 3\xi + 3C\|f\|_{lip}3\epsilon^{-1}n^{-\frac{1}{2\gamma}} + 8\epsilon \text{var}_1(L^n f) \\ &\leq 3\xi + 3C\|f\|_{lip}3\epsilon^{-1}n^{-\frac{1}{2\gamma}} + 8\epsilon[\lambda^n \alpha^n \text{var}_1(f) + B(1 + \sqrt{2}\epsilon_2^{-1}\epsilon^{-1} + 1)] \\ &\leq 3\xi + 3C\|f\|_{lip}3\epsilon^{-1}n^{-\frac{1}{2\gamma}} + 16\epsilon[\lambda^n \alpha^n \text{var}_1(f) + B(1 + \sqrt{2}\epsilon_2^{-1}\epsilon^{-1} + 1)] \\ &\leq 3\xi + 3C\sqrt{2}\epsilon_2^{-1}\epsilon^{-1}3\epsilon^{-1}n^{-\frac{1}{2\gamma}} + 16\epsilon[\lambda^n \alpha^n \text{var}_1(f) + B(1 + \sqrt{2}\epsilon_2^{-1}\epsilon^{-1} + 1)] \end{aligned}$$

Taking $\xi = n^{-\frac{1}{2\gamma}}$, $\epsilon_2 = n^{-\frac{1}{8\gamma}}$, $\epsilon = n^{-\frac{1}{8\gamma}}$ recalling that $\text{var}_1(f) \leq \text{var}_1(\nu) \leq 1$

$$\begin{aligned} \|L^n(fm)\|_1 &\leq 3n^{-\frac{1}{2\gamma}} + 9C\sqrt{2}n^{\frac{3}{8\gamma}}n^{-\frac{1}{2\gamma}} + 16n^{-\frac{1}{8\gamma}}(\alpha\lambda^n + 2B) + \sqrt{2}Bn^{-\frac{1}{8\gamma}} \\ &\leq C_3n^{-\frac{1}{8\gamma}}. \end{aligned}$$

Finally, by Equation 23

$$\begin{aligned} \|L^n \nu\|_1 &\leq \|L^n(\nu - fm)\|_{1^n} + \|L^n(fm)\|_{1^n} \\ &\leq 2\epsilon\|\nu\|_{1-BV} + \|L^n(fm)\|_{1^n} + 3\epsilon_2 \\ &\leq C_4n^{-\frac{1}{8\gamma}}. \end{aligned}$$

■

Once we have an estimate for the speed of convergence to equilibrium, by Proposition 21, and Remark 4, the following holds directly:

Proposition 28 *Let Consider a family of skew product maps $F_\delta = (T_\delta, G_\delta)$, $0 \leq \delta \leq D$ satisfying uniformly Sk1, ..., Sk3 and let $f_\delta \in \mathcal{L}^1$ its invariant probability measures, suppose*

1. F_0 is a piecewise constant toral extension as defined in Section 6, with linear Diophantine type $\gamma_l(\theta)$;
2. for each $\delta \leq D$, $d_S(T_0, T_\delta) \leq \delta$;
3. for each $\delta \leq D$ there is a set A_2 such that $m(A_2) \geq 1 - \delta$ and for all $x \in A_2, y \in M$: $|G_0(x, y) - G_\delta(x, y)| \leq \delta$.

Then for each $\gamma > \gamma_l(\theta)$ there is K_1 such that for δ small enough

$$\|f_\delta - f_0\|_{1^*} \leq K_1 \delta^{1 - \frac{1}{s\gamma + 1}}.$$

6.3 An example having Hölder behavior

In this section we show a simple example of perturbation of toral extensions satisfying assumptions *Te1* and *Te2* for which the statistical behavior is actually, only Hölder stable. This shows how that Theorem 21 gives a general estimate, which is quite sharp in the case of piecewise constant toral extensions.

Proposition 29 *Consider a well approximable Diophantine irrational θ with $\gamma_l(\theta) > 2$. Let us consider the map $F_0 : [0, 1] \times \mathcal{T}^1$ defined as a skew product $F_0(T_0(x), G_0(x, y))$ where*

$$T_0(x) = 2x \pmod{1}$$

and

$$G_0(x, y) = y + \theta\varphi(x)$$

where $\varphi = \chi_{[\frac{1}{2}, 1]}$. There is a sequence of reals $\delta_j \geq 0$, $\delta_j \rightarrow 0$ and a sequence of perturbed maps $\hat{F}_{\delta_j}(x, y) = (\hat{T}_{\delta_j}(x), \hat{G}_{\delta_j}(x, y))$ satisfying *Sk1*, ..., *Sk3*, with $\hat{T}_{\delta_j}(x) = T_0(x)$ and $\|\hat{G}_{\delta_j}(x, y) - G_0(x, y)\|_\infty \leq 2\delta_j$ such that for every $\gamma' > \gamma_l(\theta)$

$$\|\mu_0 - \mu_j\|_{1^*} \geq \frac{1}{9}(\delta_j)^{\frac{1}{\gamma'-1}}$$

holds for every j and every μ_j , invariant measure of $\hat{F}_{\delta_j}(x, y)$ in \mathcal{L}^1 .

Proof. We remark that since there is convergence to equilibrium for F_0 , the Lebesgue measure μ_0 on $[0, 1] \times \mathcal{T}^1$ is the unique invariant measure in \mathcal{L}^1 for F_0 . Consider $F_\delta = (T_0(x), y + (\delta + \theta)\varphi(x))$. For a sequence of δ_n converging to 0 it holds that $(\delta + \theta)$ is rational. For this sequence the map $y \rightarrow y + (\delta + \theta)$ ($: \mathcal{T}^1 \rightarrow \mathcal{T}^1$) is such that, 0 has a periodic orbit. Let $y_1 = 0, \dots, y_k$ be this orbit. For these parameters, consider the product measure $\mu_n = \frac{1}{k} \sum_{i \leq k} m \otimes \delta_{y_i}$, where m is the Lebesgue measure on $[0, 1]$ and δ_{y_i} is the delta measure placed on y_i . The measure μ_n is invariant for $F_\delta(x, y)$. and is in \mathcal{L}^1 . It is easy to see that $\|\mu_0 - \mu_n\|_{1^*} \geq \frac{1}{9} \frac{1}{k}$. Now the Diophantine type of θ will give an estimate for the relation between δ and k . Indeed let $\gamma' > \gamma(\theta)$, by the Diophantine

type of θ we know that there are infinitely many k_j and integers p_j such that $|k_j\theta - p_j| \leq |\frac{1}{k_j}|^{\gamma'}$ then $|\theta - \frac{p_j}{k_j}| \leq |\frac{1}{k_j}|^{\gamma'-1}$. Let us now consider $\delta_j = -\theta + \frac{p_j}{k_j}$, it holds $|\delta_j| \leq |\frac{1}{k_j}|^{\gamma'-1}$ and the angle $(\delta_j + \theta)$ generates a periodic orbit of period k_j . This happens by perturbing the second coordinate of the map by a quantity which is less than $|\frac{1}{k_j}|^{\gamma'-1}$. Summarizing, for the map F_{δ_j} we have that there is no perturbation on the first coordinate, for the second coordinate, $\|G_0 - G_{\delta_j}\|_{\infty} \leq \delta_j$ and denoting as μ_j the invariant measure on the periodic orbit defined before it holds

$$\|\mu_0 - \mu_j\|_1 \geq \frac{1}{9}(\delta_j)^{\frac{1}{\gamma'-1}}.$$

This example can further be improved by perturbing the map F_{δ_j} to a new map \hat{F}_{δ_j} in a way that μ_j (a measure supported on the attractor of \hat{F}_{δ_j}) and $\mu_j + \frac{k_j}{2}$ ⁵ (supported on the repeller of \hat{F}_{δ_j}) are the only invariant measures in \mathcal{L}^1 for \hat{F}_{δ_j} and μ_j is the unique physical measure for the system. This can be done by making a small further C^∞ perturbation on G . Let us denote again by y_1, \dots, y_{k_j} the periodic orbit of 0 as before. Let us consider a C^∞ function $g : [0, 1] \rightarrow [0, 1]$ such that:

- g is negative on the each interval $[y_i, y_i + \frac{1}{2k_j}]$ and positive on each interval $[y_i + \frac{1}{2k_j}, y_{i+1}]$ (so that $g(y_i + \frac{1}{2k_j}) = 0$);
- g' is positive in each interval $[y_i + \frac{1}{3k_j}, y_{i+1} - \frac{1}{3k_j}]$ and negative in $[y_i, y_{i+1}] - [y_i + \frac{1}{3k_j}, y_{i+1} - \frac{1}{3k_j}]$.

Considering $D_\delta : \mathcal{T}^1 \rightarrow \mathcal{T}^1$ defined by $D_\delta(x) = x + \delta g(x) \bmod(1)$ it holds that the iteration of this map send all the space but the set $\{y_i + \frac{1}{2k_j} \text{ s.t. } i \leq k_j\}$ (which is a repeller) to the set $\{y_i \text{ s.t. } i \leq k_j\}$ (the attractor). Then define \hat{F}_{δ_j} as:

$$\hat{F}_{\delta_j}(x, y) = (T_{\delta_j}(x), D_{\delta_j}(y + (\delta_j + \theta)\varphi(x))).$$

The claim directly follows from the remark that for the map $(\hat{F}_{\delta_j})^{k_j}$ the sets $\Gamma_1 := [0, 1] \times \{y_i \text{ s.t. } i \leq k_j\}$ and $\Gamma_2 := [0, 1] \times \{y_i + \frac{1}{2k_j} \text{ s.t. } i \leq k_j\}$ are invariant and the set Γ_1 attracts the whole $[0, 1] \times \mathcal{T}^1 - \Gamma_2$. ■

The construction done in the previous proof can be extended to show Hölder behavior for the average of a given regular observable. We show an explicit example of such an observable for a skew product with a particular angle θ .

Proposition 30 *Consider a map F_0 as above with the rotation angle $\theta = \sum_{i=1}^{\infty} 2^{-2^i}$. With*

$$T_0(x) = 2x \bmod(1)$$

⁵Defined as $[\mu_j + \frac{1}{2k_j}](A) = \mu_j(A - \frac{1}{2k_j})$ for each measurable set A in \mathcal{T}^1 . Where $A - \frac{1}{2k_j}$ is the translation of the set A by $-\frac{1}{2k_j}$.

and

$$G_0(x, y) = y + \theta\varphi(x)$$

as in Proposition 29. Let \hat{F}_{δ_j} be its perturbations as described in the proof of the proposition and μ_j their invariant measures in \mathcal{L}^1 . There is an observable $\psi : [0, 1] \times \mathcal{T}^1 \rightarrow \mathbb{R}$ with derivative in L^2 and $C \geq 0$ such that

$$|\int \psi d\mu_0 - \int \psi d\mu_j| \geq C\sqrt{\delta_j}.$$

Proof. We recall that $\sum_{n+1}^{\infty} 2^{-2^{2^i}} \leq 2^{-2^{2(n+1)+1}}$. By this $\|2^{2^{2n}}\theta\| \leq 2^{-2^{2(n+1)+1}}$ and the Diophantine type of θ is greater than 4. Following the construction above, we have that with a perturbation of size less than $2^{-2^{2(n+1)+1}}$ the angles $\theta_j = \sum_{i=1}^j 2^{-2^{2^i}}$ generate on the second coordinate of the skew product orbits of period $2^{2^{2j}}$. Now let us construct a suitable observable which can "see" the change of the invariant measure under this perturbation. Let us consider

$$\psi(x, y) = \sum_{i=1}^{\infty} \frac{1}{(2^{2^{2^i}})^2} \cos(2^{2^{2^i}} 2\pi y) \quad (25)$$

and $\psi_k(x, y) = \sum_{i=1}^k \frac{1}{(2^{2^{2^i}})^2} \cos(2^{2^{2^i}} 2\pi y)$ Since for the observable ψ , the i -th Fourier coefficient decreases like i^{-2} , then ψ has a derivative in L^2 . Let $x_1 = 0, \dots, x_{2^{2^{2j}}}$ be the periodic orbit of 0 for $y \rightarrow y + \theta_j$, and $\mu_j = \frac{1}{2^{2^{2j}}} \sum \delta_{x_i}$ the physical measure supported on it. Since $2^{2^{2j}}$ divides $2^{2^{2(j+1)}}$ then $\sum_{i=1}^{2^{2^{2j}}} \psi_k(x_i) = 0$ for every $k < j$, thus $\int \psi_{j-1} d\mu_j = 0$. Then

$$\begin{aligned} v_j &= \int \psi d\mu_j \geq \frac{1}{(2^{2^{2j}})^2} - \sum_{i=j+1}^{\infty} \frac{1}{(2^{2^{2^i}})^2} \\ &\geq 2^{-2^{2j+1}} - 2^{-2^{2(j+1)+1}}. \end{aligned}$$

And for j big enough

$$2^{-2^{2j+1}} - 2^{-2^{2(j+1)+1}} \geq \frac{1}{2}(2^{-2^{2j}})^2.$$

Summarizing, with a perturbation of size $\delta_j = \sum_{j+1}^{\infty} 2^{-2^{2^i}} \leq 2 * 2^{-2^{2(j+1)}} = 2^{-2^{2(j+1)}} = 2(2^{-2^{2j}})^4$ we get a change of average for the observable ψ from $\int \psi dm = 0$ to $v_n \geq \frac{1}{2}(2^{-2^{2j}})^2$. Hence we have that there is a C such that with a perturbation of size δ_j we get a change of average for the observable ψ of size bigger than $C\sqrt{\delta_j}$. ■

Remark 31 Using $\frac{1}{(2^{2^{2^i}})^{\alpha}}$ instead of $\frac{1}{(2^{2^{2^i}})^2}$ in (25) we can obtain a smoother observable. Using rotation angles with bigger and bigger Diophantine type it is possible to obtain a dependence of the physical measure to perturbations with worse and worse Hölder exponent. Using angles with infinite Diophantine type it is possible to have a behavior whose modulus of continuity is worse than the Hölder one.

References

- [1] J. Alves, *Strong statistical stability of non-uniformly expanding maps*. Nonlinearity, 17, 4, 1193-1215 (2004).
- [2] J. Alves, M. Viana *Statistical stability for robust classes of maps with non-uniform expansion*. Ergodic Theory and Dynam. Systems, 22 , 1, 1-32 (2002).
- [3] J. F. Alves, M. Soufi *Statistical Stability in Chaotic Dynamics* Progress and Challenges in Dyn. Sys. Springer Proc. in Math. & Statistics V. 54, 2013, pp 7-24
- [4] V. Araújo, S. Galatolo, M.J, Pacifico *Decay of correlations for maps with uniformly contracting fibers and logarithm law for singular hyperbolic attractors*. Math. Z. 276 (2014), no. 3-4, 1001–1048.
- [5] W. Bahsoun, S. Vaienti *Metastability of certain intermittent maps* Nonlinearity 25, 1, 107, (2012)
- [6] V. Baladi., L-S Young, *On the spectra of randomly perturbed expanding maps* Comm. Math. Phys. 156 , 355-385 (1993)
- [7] V. Baladi *Linear response, or else* ICM Seoul 2014 talk (arXiv:1408.2937)
- [8] V. Baladi, M. Todd *Linear response for intermittent maps* arXiv:1508.02700
- [9] V. Baladi , M. Benedicks , N. Schnellmann *Whitney-Holder continuity of the SRB measure for transversal families of smooth unimodal maps* Invent. Math. 201 773-844 (2015)
- [10] V. Baladi, T. Kuna and V. Lucarini *Linear and fractional response for the SRB measure of smooth hyperbolic attractors and discontinuous observables* Nonlinearity 30 1204-1220 (2017)
- [11] W. Bahsoun, B. Saussol *Linear response in the intermittent family: differentiation in a weighted C^0 -norm*. arXiv:1512.01080
- [12] A. Boyarsky, P. Gora *Laws of Chaos - Invariant Measures and Dynamical Systems in One Dimension*. Birkhauser Boston, 1997.
- [13] M. I. Brin. *The topology of group extensions of C-systems*. Mat. Zametki, 18(3):453–465, 1975.
- [14] K. Burns A. Wilkinson. *Stable ergodicity of skew products*. Ann. Sci. ´ Ecole Norm. Sup. (4), 32(6):859–889, 1999.
- [15] O Butterley P. Eslami *Exponential Mixing for Skew Products with Discontinuities* . Trans. Amer. Math. Soc., 369, 783–803, 2017.

- [16] D. Dolgopyat. *On dynamics of mostly contracting diffeomorphisms* , Comm. in Math. Physics , 213 (2000) 181-201.
- [17] D. Dolgopyat. *On mixing properties of compact group extensions of hyperbolic systems*. Israel J. Math., 130:157–205, 2002.
- [18] D. Dolgopyat *On differentiability of SRB states for partially hyperbolic systems* , Invent. Math. 155 (2004) 389–449.
- [19] S. Galatolo *Quantitative statistical stability and convergence to equilibrium. An application to maps with indifferent fixed points* arXiv:1512.01646
- [20] S. Galatolo, R. Lucena, *Spectral Gap and quantitative statistical stability for systems with contracting fibers and Lorenz like maps*. arXiv:1507.08191
- [21] S. Galatolo, J Rousseau , B Saussol *Skew products, quantitative recurrence, shrinking targets and decay of correlations* Erg. Th. Dyn. Sys 35, 6 (2015), pp. 1814-1845
- [22] S. Galatolo, Nisoli I., *An elementary approach to rigorous approximation of invariant measures*. *SIAM J. Appl. Dyn. Syst.* 13 (2014), no. 2, 958–985.
- [23] S Galatolo, M Monge, I Nisoli *Rigorous approximation of stationary measures and convergence to equilibrium for iterated function systems* J. Phys. A: Math. and Theor. 49 (27), 274001
- [24] S. Galatolo and P. Peterlongo *Long hitting time, slow decay of correlations and arithmetical properties* Disc. Cont. Dyn. Sys. 27,n 1, (2010) pp 185 - 204
- [25] A. Korepanov *Linear response for intermittent maps with summable and nonsummable decay of correlations*. arXiv:1508.06571
- [26] G. Keller. *Generalized bounded variation and applications to piecewise monotonic transformations* Z. Wahrsch. Verw. Gebiete, 69(3):461–478, (1985).
- [27] G. Keller, C. Liverani *Stability of the spectrum for transfer operators* Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 28 no. 1, 141-152 (1999).
- [28] C. Liverani *Invariant measures and their properties. A functional analytic point of view*, Dynamical Systems. Part II: Topological Geometrical and Ergodic Properties of Dynamics. Proceedings, by Scuola Normale Superiore, Pisa (2004).
- [29] F. Naud *On the rate of mixing of circle extensions of Anosov maps*. arXiv:1612.05011
- [30] F. Naud *Exponentially mixing, locally constant skew extensions of shift maps* arXiv:1701.08659

- [31] D. Ruelle. *Flots qui ne m'élagent pas exponentiellement*. C. R. Acad. Sci. Paris. I Math., 296(4):191–193, 1983.
- [32] W. Shen, S. Van Strien *On stochastic stability of expanding circle maps with neutral fixed points* Dynamical Systems: An International Journal 28, 3, 423-452 (2013)
- [33] Z. Zhang *On the smooth dependence of SRB measures for partially hyperbolic systems* arXiv:1701.05253